Lecture 6–Numerical integration

Instructor: Dr. C.J. Xie (macjxie@ust.hk)

1 Numerical integration

Notations

- L_n denotes the left rectangular sum approximation using n rectangles;
- R_n denotes the right rectangular sum approximation using n rectangles;
- M_n denotes the midpoint rectangular sum approximation using n rectangles;
- T_n denotes the trapezoid approximation using n trapezoids;
- S_n denotes the Riemann sum by Simpson's rule.

Basic ideas

When all methods we have learned until now do not work, we need to consider how to calculate the integral numerically by the definition.
Definition of definite integral

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*)\Delta x,$$

where $\Delta x = \frac{b-a}{n}$, x_i^* is taken in the subinterval.
Rk. The integration can be approximated by the Riemann sum,

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{n} f(x_i^*)\Delta x.$$

More accuracy of approximation as n larger.
Example. Evaluate $\int_0^1 e^{x^2}\,dx$.
solution. Note that we could not use the Fundamental Theorem of Calculus directly, since $F'(x) = e^{x^2}$, where F has not explicit expression in terms of the elementary functions. The method of integration by parts or any substitutions can not work as well.
By the definition of definite integral, let $f(x) = e^{x^2}$, the domain $[0,1]$, thus, the length of subintervals $\Delta x = \frac{1-0}{n} = \frac{1}{n}$, the subintervals are then

$$[0,\Delta x]\,,\ [\Delta x, 2\Delta x]\,,\cdots,\ [(n-2)\Delta x,(n-1)\Delta x]\,,\ [(n-1)\Delta x, n\Delta x]\,,$$

say,

$$\left[0,\frac{1}{n}\right],\ \left[\frac{1}{n},2\cdot\frac{1}{n}\right],\cdots,\ \left[(n-2)\cdot\frac{1}{n},(n-1)\cdot\frac{1}{n}\right],\ \left[(n-1)\cdot\frac{1}{n},n\cdot\frac{1}{n}\right],$$

Using the rectangular domain under the graph of f on each subinterval, we have

- if taking $x_i^L = (i-1)\cdot\frac{1}{n}$, $i = 1, 2\cdots, n$, to be **the left end points**, the Riemann sum is

$$L_n = \sum_{i=1}^{n} f(x_i^L)\cdot\frac{1}{n} = \sum_{i=1}^{n} e^{(x_i^L)^2}\cdot\frac{1}{n},$$

- if taking $x_i^R = i\cdot\frac{1}{n}$, $i = 1, 2\cdots, n$, to be **the right end points**, the Riemann sum is

$$R_n = \sum_{i=1}^{n} f(x_i^R)\cdot\frac{1}{n} = \sum_{i=1}^{n} e^{(x_i^R)^2}\cdot\frac{1}{n},$$

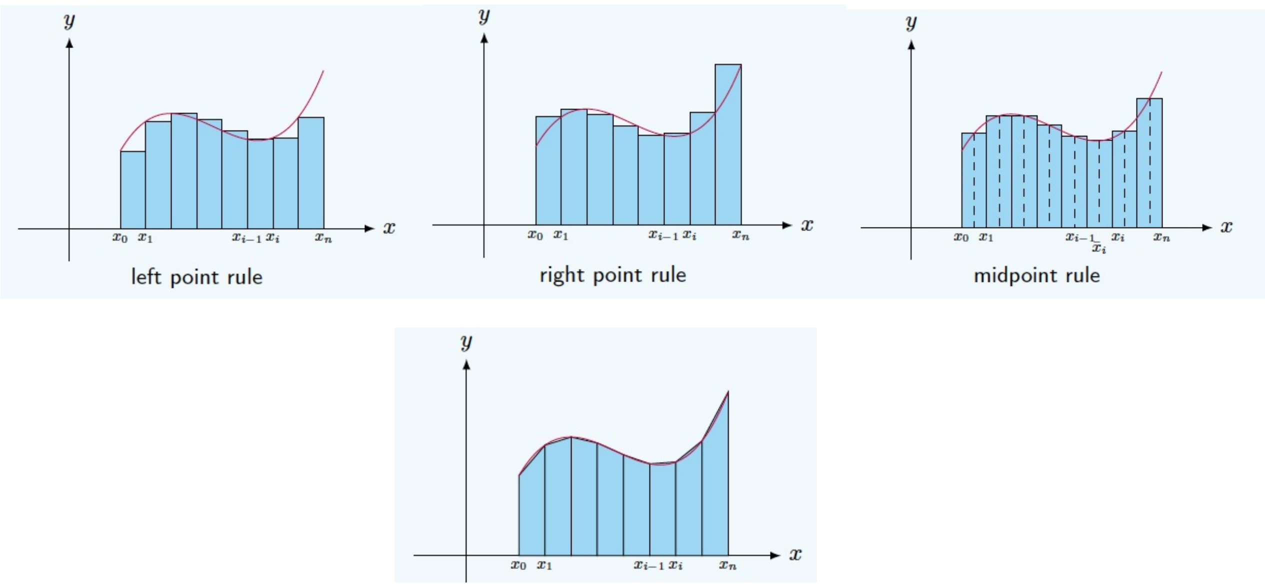
- if taking $x_i^M = \left(\frac{x_i^L + x_i^R}{2}\right)\cdot\frac{1}{n} = \left(\frac{i+i-1}{2}\right)\cdot\frac{1}{n}$, $i = 1, 2\cdots, n$, to be **the middle points**, the Riemann sum is

$$M_n = \sum_{i=1}^{n} f(x_i^M)\cdot\frac{1}{n} = \sum_{i=1}^{n} e^{(x_i^M)^2}\cdot\frac{1}{n},$$

or using the trapezoidal domain under the graph of f on each subinterval, we have

$$T_n = \sum_{i=1}^{n}\left[\frac{f(x_i^L)+f(x_i^R)}{2}\right]\cdot\frac{1}{n} = \sum_{i=1}^{n}\left[\frac{e^{(x_i^L)^2}+e^{(x_i^R)^2}}{2}\right]\cdot\frac{1}{n},$$

See the geometric interpretations of each rule below in Figure 1.

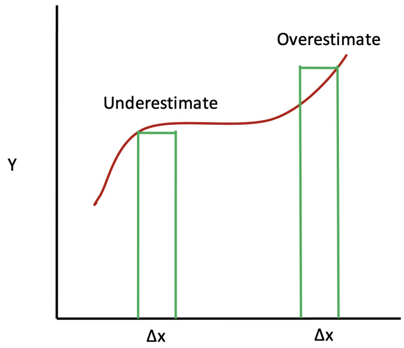


(d) tapezoidal rule

Figure 1: Different rules for the numerical integration.

Rk the overestimate and underestimate:

Method	Overestimates	Underestimates
Left L_n	$f'(x) < 0$ (decreasing)	$f'(x) > 0$ (increasing)
Right R_n	$f'(x) > 0$ (increasing)	$f'(x) < 0$ (decreasing)
Midpoint M_n	$f''(x) < 0$	$f''(x) > 0$
Trapezoidal T_n	$f''(x) > 0$	$f''(x) < 0$
Simpson S_n	$f^{(4)}(x) > 0$	$f^{(4)}(x) < 0$



Example (from classviva.org). Estimate $\int_0^1 \cos x^2\, dx$ using the (a) Trapezoidal rule and (b) middle point rule, each with $n = 4$.

solution

Let $f(x) = \cos x^2$, $\int_0^1 f(x)\, dx \approx \sum_{i=1}^n f(x_i^*)\Delta x$, where $\Delta x = \frac{1-0}{4} = \frac{1}{4}$, the subintervals are then

$$\left[0, \frac{1}{4}\right], \left[\frac{1}{4}, 2\cdot\frac{1}{4}\right], \left[2\cdot\frac{1}{4}, 3\cdot\frac{1}{4}\right], \left[3\cdot\frac{1}{4}, 4\cdot\frac{1}{4}\right]$$

Thus, the **left points** are $x_i^L = (i-1)\cdot\frac{1}{4}$, $i = 1,2,3,4$ and the **left points** are $x_i^R = i\cdot\frac{1}{4}$, $i = 1,2,3,4$.

(a) by the trapezoidal rule, the Riemann sum is given by

$$\begin{aligned} T_4 &= \sum_{i=1}^4 \left[\frac{f(x_i^L) + f(x_i^R)}{2}\right]\cdot\frac{1}{4} = \sum_{i=1}^4 \left[\frac{\cos[(i-1)\cdot\frac{1}{4}]^2 + \cos[i\cdot\frac{1}{4}]^2}{2}\right]\cdot\frac{1}{4} \\ &= \frac{1}{8}\left[\cos 0 + 2\cos\frac{1}{16} + 2\cos\frac{1}{4} + 2\cos\frac{9}{16} + \cos 1\right] \approx 0.89576, \end{aligned}$$

(b) by the middle point rule, the Riemann sum is given by

$$\begin{aligned} M_4 &= \sum_{i=1}^4 f\left(\frac{x_i^L + x_i^R}{2}\right)\cdot\frac{1}{4} = \sum_{i=1}^4 \cos\left(\frac{\frac{i-1}{4} + \frac{i}{4}}{2}\right)^2\cdot\frac{1}{4} \\ &= \frac{1}{4}\left[\cos\frac{1}{64} + \cos\frac{9}{64} + \cos\frac{25}{64} + \cos\frac{49}{64}\right] \approx 0.90891 \end{aligned}$$

Error evaluation

Example. Estimate $\mathcal{I} = \int_0^1 \frac{1}{1+x^2}\, dx$ by the left point rule, the right point rule, the midpoint rule or the trapezoidal rule (usually to (five or) six decimal places).

solution

The exact value is

$$\mathcal{I} = \int_0^1 \frac{1}{1+x^2}\, dx = \tan^{-1} x|_0^1 = \tan^{-1} 1 - \tan^{-1} 0 = \frac{\pi}{4} - 0 = \frac{\pi}{4} \approx 0.785398.$$

We take $n = 4$ equal subintervals and $f(x) = \frac{1}{1+x^2}$, $\int_0^1 f(x)\, dx \approx \sum_{i=1}^n f(x_i^*)\Delta x$, where the length $\frac{1-0}{4}$, the subintervals are then

$$\left[0, \frac{1}{4}\right], \left[\frac{1}{4}, 2\cdot\frac{1}{4}\right], \left[2\cdot\frac{1}{4}, 3\cdot\frac{1}{4}\right], \left[3\cdot\frac{1}{4}, 4\cdot\frac{1}{4}\right]$$

Thus, the **left points** are $x_i^L = (i-1)\cdot\frac{1}{4}$, $i = 1,2,3,4$ and the **right points** are $x_i^R = i\cdot\frac{1}{4}$, $i = 1,2,3,4$.

- by the **left point rule**, the Riemann sum is given by

$$L_4 = \sum_{i=1}^{4} f(x_i^L) \cdot \frac{1}{4} = \sum_{i=1}^{4} \frac{1}{1+(x_i^L)^2} \cdot \frac{1}{4} = \sum_{i=1}^{4} \frac{1}{1+\left(\frac{i-1}{4}\right)^2} \cdot \frac{1}{4}$$
$$= \left[1 + \frac{16}{17} + \frac{4}{5} + \frac{16}{25}\right] \cdot \frac{1}{4} = 0.845294$$

The error is $E_{L_4} = \mathcal{I} - L_4 = -0.059896$, which indicates $\mathcal{I} < L_4$, thus, this is **over-estimate**.

- by the **right point rule**, the Riemann sum is given by

$$R_4 = \sum_{i=1}^{4} f(x_i^R) \cdot \frac{1}{4} = \sum_{i=1}^{4} \frac{1}{1+(x_i^R)^2} \cdot \frac{1}{4} = \sum_{i=1}^{4} \frac{1}{1+\left(\frac{i}{4}\right)^2} \cdot \frac{1}{4}$$
$$= \left[\frac{16}{17} + \frac{4}{5} + \frac{16}{25} + \frac{1}{2}\right] \cdot \frac{1}{4} \approx 0.720294$$

The error is $E_{R_4} = \mathcal{I} - R_4 = 0.065104$, which indicates $\mathcal{I} > R_4$, thus, this is **under-estimate**.

- by the **midpoint rule**, the Riemann sum is given by

$$M_4 = \sum_{i=1}^{4} f\left(\frac{x_i^L + x_i^R}{2}\right) \cdot \frac{1}{4} = \sum_{i=1}^{4} \frac{1}{1+\left(\frac{x_i^L+x_i^R}{2}\right)^2} \cdot \frac{1}{4} = \sum_{i=1}^{4} \frac{1}{1+\left(\frac{2i-1}{8}\right)^2} \cdot \frac{1}{4}$$
$$= \left[\frac{64}{65} + \frac{64}{73} + \frac{64}{89} + \frac{64}{113}\right] \cdot \frac{1}{4} \approx 0.786700$$

The error is $E_{M_4} = \mathcal{I} - M_4 = -0.001302$, which indicates $\mathcal{I} < M_4$, thus, this is **over-estimate**.

- by the trapezoidal rule, the Riemann sum is given by

$$T_4 = \sum_{i=1}^{4} \left(\frac{f(x_i^L) + f(x_i^R)}{2}\right) \cdot \frac{1}{4} = \sum_{i=1}^{4} \left(\frac{\frac{1}{1+(x_i^L)^2} + \frac{1}{1+(x_i^R)^2}}{2}\right) \cdot \frac{1}{4}$$
$$= \sum_{i=1}^{4} \left(\frac{\frac{1}{1+(\frac{i-1}{4})^2} + \frac{1}{1+(\frac{i}{4})^2}}{2}\right) \cdot \frac{1}{4} = \left[\frac{1+\frac{16}{17}}{2} + \frac{\frac{16}{17}+\frac{4}{5}}{2} + \frac{\frac{4}{5}+\frac{16}{25}}{2} + \frac{\frac{16}{25}+\frac{1}{2}}{2}\right] \cdot \frac{1}{4}$$
$$= \frac{L_4 + R_4}{2} \approx 0.782794$$

The error is $E_{T_4} = \mathcal{I} - T_4 = 0.002604$, which indicates $\mathcal{I} > T_4$, thus, this is **under-estimate**.

Note that absolute error $|E_{M_4}| < |E_{T_4}| < |E_{L_4}| < |E_{R_4}|$. Thus, the accuracy from high to low: **midpoint rule** $>$ **trapezoidal rule** $>$ **left point rule** $>$ **right point rule**.

Rk.

- ✓ For all rules (whatever left point rule, right point rule, midpoint rule or trapezoidal rule) produce better approximation as n increases. When $n \to \infty$, the Riemann sum is exactly the definite integral;
- ✓ Generally, $|E_{T_n}| \approx 2|E_{M_n}|$.

Error bounds

Theorem. Suppose f' is continuous and $|f'| \le K$ on $[a, b]$, and denote the length $h = \frac{b-a}{n}$, then

- for the **right point rule**,

$$|E_{R_n}| = \left| \int_a^b f(x)\,dx - R_n \right| \le \frac{K(b-a)^2}{2n} = \frac{K(b-a)}{2} h$$

- for the **left point rule**,

$$|E_{L_n}| = \left| \int_a^b f(x)\,dx - L_n \right| \le \frac{K(b-a)^2}{2n} = \frac{K(b-a)}{2} h$$

Theorem. Suppose f'' is continuous and $|f''| \le K$ on $[a, b]$ and denote the length $h = \frac{b-a}{n}$, then

- for the **trapezoidal rule**,

$$|E_{T_n}| = \left| \int_a^b f(x)\,dx - T_n \right| \le \frac{K(b-a)^3}{12n^2} = \frac{K(b-a)}{12} h^2,$$

- for the **midpoint rule**,

$$|E_{M_n}| = \left| \int_a^b f(x)\,dx - M_n \right| \le \frac{K(b-a)^3}{24n^2} = \frac{K(b-a)}{24} h^2.$$

Example. Consider $\mathcal{I} = \int_1^5 \cos\left(\frac{1}{x}\right)$. How large does n need to be to use the **left point rule**, **right point rule**, **midpoint rule** and **trapezoidal rule** to approximate $\mathcal{I}$ with error guaranteed to be less than the tolerance 10^{-5}.

solution The length $h = \frac{5-1}{n} = \frac{4}{n}$. Let $f = \cos\left(\frac{1}{x}\right)$, thus, $|f'| = |\sin\left(\frac{1}{x}\right) \cdot \left(-\frac{1}{x^2}\right)| \le |\frac{1}{x^2}|$. Note that $1 \le x \le 5$, thus, $1 \le x^2 \le 25$, and $\frac{1}{25} \le \frac{1}{x^2} \le 1$. So $|f'| \le 1$. Taking $K = 1$.

- for the **left or right point rule**, let

$$\frac{K(b-a)}{2} h = 2 \cdot \frac{4}{n} = \frac{8}{n} < 10^{-5} \Longrightarrow n > \frac{8}{10^{-5}} = 8 \times 10^5,$$

  thus, $n = 8 \times 10^5 + 1 = 800001$,

- for the **midpoint rule**, let

$$\frac{K(b-a)}{24} h^2 = \frac{1}{6} \cdot \frac{16}{n^2} = \frac{8}{3n^2} < 10^{-5} \Longrightarrow n > \sqrt{\frac{8}{3 \times 10^{-5}}} \approx 516.40$$

  thus $n = 517$,

- for the **trapezoidal rule**, let

$$\frac{K(b-a)}{12} h^2 = \frac{1}{3} \cdot \frac{16}{n^2} = \frac{16}{3n^2} \le 10^{-5} \Longrightarrow n > \sqrt{\frac{16}{3 \times 10^{-5}}} \approx 730.30,$$

  thus, $n = 731$.

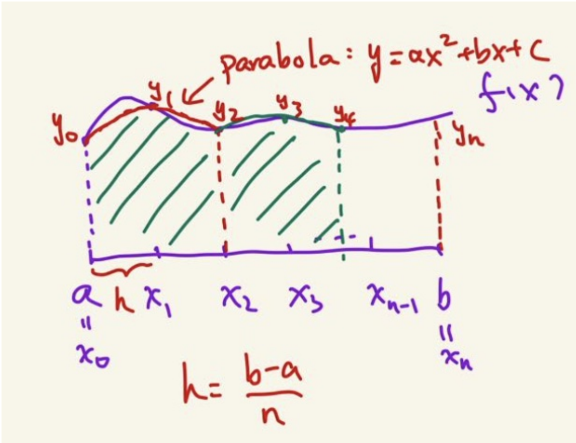


Figure 2: If we have n subintervals which are $[x_0, x_1], [x_1, x_2], [x_2, x_3], \cdots, [x_{n-1}, x_n]$, where n is even number. For $[x_0, x_2]$, the area of parabola $y = ax^2+bx+c$ passes the three points $(x_0, y_0))$, (x_1, y_1), (x_2, y_2), the base length $h = \frac{b-a}{n}$. For $[x_2, x_4], \cdots, [x_{n-2}, x_n]$, the parabola can be constructed as well.

Simpson's rule

We have use the area shape for different rules below,

- the rectangular area for **left point rule**, **the right rule** and the **midpoint rule**;
- the trapezoids for the **trapezoidal rule**;

These rules are based on the Riemann sum over each subintervals. In each subinterval, the area of the shape is constructed by two points information, see Figure 1.
Now we construct parabolas below for each group of **three points** in Figure 2.
Assume the parabola function $f = ax^2 + bx + c$, passing the three points (x_0, y_0), (x_1, y_1), (x_2, y_2), $h = \frac{b-a}{n}$ with n being the even number of partition. Since the parabola through the three points, we have

$$\begin{aligned} y_0 &= f(x_0) = a(x_0)^2 + bx_0 + c, \\ y_1 &= f(x_1) = a(x_1)^2 + bx_1 + c, \\ y_2 &= f(x_2) = a(x_2)^2 + bx_2 + c. \end{aligned}$$

On the other hand, the area under the parabola is given by

$$\begin{aligned} A_1 &= \int_{x_0}^{x_2} (ax^2 + bx + c)\, dx = \left(\frac{1}{3}ax^3 + \frac{1}{2}bx^2 + cx \right) \Big|_{x_0}^{x_2} \\ &= \frac{1}{3}a(x_2)^3 + \frac{1}{2}bx_2^2 + cx_2 - \frac{1}{3}a(x_0)^3 - \frac{1}{2}bx_0^2 - cx_0 \\ &= \frac{1}{3}a(x_2 - x_0)\left[(x_2)^2 + x_2x_0 + (x_0)^2\right] + \frac{1}{2}b(x_2 - x_0)(x_2 + x_0) + c(x_2 - x_0) \\ &= \frac{1}{6}(x_2 - x_0)\left[2a(x_2)^2 + 2ax_2x_0 + 2a(x_0)^2 + 3b(x_2 + x_0) + 6c\right] \\ &= \frac{1}{6} \cdot 2h \left[(ax_0^2 + bx_0 + c) + 4\left(a\left[\frac{x_0 + x_2}{2}\right]^2 + b\left[\frac{x_0 + x_2}{2} + c\right] \right) + (ax_2^2 + bx_2 + c) \right] \\ &= \frac{1}{3}h\left[f(x_0) + 4f(x_1) + f(x_2)\right] \end{aligned}$$

where $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$, $x_2 - x_1 = 2h$, $x_1 = \frac{x_0+x_2}{2}$ has been applied. By the same way, we have the parabola area for next group of three points below,

$$A_2 = \frac{1}{3}h\left[f(x_2) + 4f(x_3) + f(x_4)\right].$$

So on and so forth, for the interval $[x_{2i-1}, x_{2i}]$ we have

$$A_i = \frac{1}{3}h\left[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})\right].$$

and for the interval $[x_{n-2}, x_n]$

$$A_{\frac{n}{2}} = \frac{1}{3}h\left[f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\right].$$

Thus, for $h = \frac{b-a}{n}$, we have

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{\frac{n}{2}} A_i = \sum_{i=1}^{\frac{n}{2}} \frac{1}{3}h\left[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})\right] := S_n,$$

which is called **Simpson's rule**.
Especially,

- if taking $n = 2$, we have $h = \frac{b-a}{2}$, and

$$S_2 = \frac{1}{3}h\left[1f(x_0) + 4f(x_1) + 1f(x_2)\right],$$

- if taking $n = 4$, we have $h = \frac{b-a}{4}$, and

$$\begin{aligned} S_4 &= \sum_{i=1}^{2} \frac{1}{3}h\left[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})\right], \\ &= \frac{1}{3}h\left[f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4)\right] \\ &= \frac{1}{3}h\left[1f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 1f(x_4)\right], \end{aligned}$$

- if taking $n = 6$, we have $h = \frac{b-a}{6}$, and

$$\begin{aligned} S_6 &= \sum_{i=1}^{3} \frac{1}{3}h\left[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})\right], \\ &= \frac{1}{3}h\left[f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4) + f(x_4) + 4f(x_5) + f(x_6)\right] \\ &= \frac{1}{3}h\left[1f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + 4f(x_5) + 1f(x_6)\right], \end{aligned}$$

- if taking $n = 8$, we have $h = \frac{b-a}{8}$, and

$$\begin{aligned} S_8 &= \sum_{i=1}^{4} \frac{1}{3}h\left[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})\right], \\ &= \frac{1}{3}h\big[f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4) \\ &\quad + f(x_4) + 4f(x_5) + f(x_6) + f(x_6) + 4f(x_7) + f(x_8)\big] \\ &= \frac{1}{3}h\left[1f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + 4f(x_5) + 2f(x_6) + 4f(x_7) + 1f(x_8)\right], \end{aligned}$$

Example. Estimate $\int_0^1 \cos x^2\,dx$ using the **Simpson's rule** with $n = 4$.
solution.

Let $f(x)=\cos x^2$, the base length $h=\Delta x=\frac{1-0}{4}$, the subintervals are then

$$[0,\Delta x]\,,\quad [\Delta x, 2\cdot\Delta x]\,,\quad [2\cdot\Delta x, 3\cdot\Delta x]\,,\quad [3\cdot\Delta x, 1]\,,$$

say,

$$\left[0,\frac{1}{4}\right],\quad \left[\frac{1}{4},2\cdot\frac{1}{4}\right],\quad \left[2\cdot\frac{1}{4},3\cdot\frac{1}{4}\right],\quad \left[3\cdot\frac{1}{4},1\right].$$

By Simpson's rule, we have

$$\begin{aligned}
S_4 &= \sum_{i=1}^{\frac{4}{2}} \frac{1}{3}h\left[f(x_{2i-2})+4f(x_{2i-1})+f(x_{2i})\right]\\
&= \frac{1}{3}h\left[f(x_0)+4f(x_1)+f(x_2)+f(x_2)+4f(x_3)+f(x_4)\right]\\
&= \frac{1}{3}h\left[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+f(x_4)\right]\\
&= \frac{1}{3}\cdot\frac{1}{4}\left[\cos 0+4\cos\left(\frac{1}{4}\right)^2+2\cos\left(\frac{1}{2}\right)^2+4\cos\left(\frac{3}{4}\right)^2+\cos 1\right]\\
&\approx 0.90450.
\end{aligned}$$

Example (evaluation the error). Estimate $\mathcal{I}=\int_0^1 \frac{1}{1+x^2}\,dx$ by the **Simpson's rule** with $n=4$ (to six decimal places).

solution. This example has been evaluated by **left point rule**, **right point rule**, **midpoint rule** and **trapezoidal rule** before. Note that the exact value is given by the FTC, $\mathcal{I}=\frac{\pi}{4}\approx 0.785398$. By the **Simpson's rule**, we have the base length $h=\Delta x=\frac{1-0}{4}$, and

$$\begin{aligned}
S_4 &= \sum_{i=1}^{2} \frac{1}{3}h\left[f(x_{2i-2})+4f(x_{2i-1})+f(x_{2i})\right],\\
&= \frac{1}{3}h\left[f(x_0)+4f(x_1)+f(x_2)+f(x_2)+4f(x_3)+f(x_4)\right]\\
&= \frac{1}{3}h\left[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+f(x_4)\right]\\
&= \frac{1}{3}\cdot\frac{1}{4}\left[\frac{1}{1+0}+4\cdot\frac{1}{1+(\frac{1}{4})^2}+2\cdot\frac{1}{1+(\frac{1}{2})^2}+4\cdot\frac{1}{1+(\frac{3}{4})^2}+\frac{1}{1+1}\right]\\
&\approx 0.785392.
\end{aligned}$$

The error is

$$E_{S_4}=\mathcal{I}-S_4=0.000006>0.$$

Thus, $\mathcal{I}>S_4 \implies$ **underestimate**. Comparing with the previous results below,

$$\begin{aligned}
|E_{L_4}| &= 0.059896\\
|E_{R_4}| &= 0.065104\\
|E_{M_4}| &= 0.001302\\
|E_{T_4}| &= 0.002604.
\end{aligned}$$

Thus, we have $|E_{S_4}|<|E_{M_4}|<|E_{T_4}|<|E_{L_4}|<|E_{R_4}|$. In general, the error of **Simpson's rule** is smaller than the error in the **trapezoidal rule** and **midpoint rule**. In this sense, **Simpson's rule** is more accurate rule to approximate the integration.

Also, Note that $T_4 \approx 0.782794$ and $M_4 \approx 0.786700$, we have

$$\frac{1}{3}T_4 + \frac{2}{3}M_4 \approx 0.785398,$$

when we take $n = 8$, we have $h = \frac{1-0}{8}$

$$\begin{aligned}
S_8 &= \sum_{i=1}^{4} \frac{1}{3}h\left[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})\right], \\
&= \frac{1}{3}h\big[f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4) \\
&\quad + f(x_4) + 4f(x_5) + f(x_6) + f(x_6) + 4f(x_7) + f(x_8)\big] \\
&= \frac{1}{3}h\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + 4f(x_5) + 2f(x_6) + 4f(x_7) + f(x_8)\right], \\
&= \frac{1}{3}\cdot\frac{1}{8}\Big[\frac{1}{1} + 4\cdot\frac{1}{1+(\frac{1}{8})^2} + 2\cdot\frac{1}{1+(\frac{2}{8})^2} + 4\cdot\frac{1}{1+(\frac{3}{8})^2} + 2\cdot\frac{1}{1+(\frac{4}{8})^2} \\
&\quad + 4\cdot\frac{1}{1+(\frac{5}{8})^2} + 2\cdot\frac{1}{1+(\frac{6}{8})^2} + 4\cdot\frac{1}{1+(\frac{7}{8})^2} + \frac{1}{1+1}\Big] \\
&\approx 0.785398
\end{aligned}$$

Thus, we have

$$\mathcal{I} \approx S_8 = \frac{1}{3}T_4 + \frac{2}{3}M_4$$

In general, we have $S_{2n} = \frac{1}{3}T_n + \frac{2}{3}M_n$, where $n = 1, 2, 3, 4, \cdots$.
Rk.

- For the **Simpson's rule**, the number of partition n should be even number, since each group of three points is used;
- $S_{2n} = \frac{1}{3}T_n + \frac{2}{3}M_n$, for example, let $n = 2$, we have

$$\begin{aligned}
S_4 &= \frac{1}{3}\cdot\frac{b-a}{4}\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + f(x_4)\right] \\
T_2 &= \sum_{i=1}^{2}\left[\frac{f(x_i^L) + f(x_i^R)}{2}\right]\frac{b-a}{2} = \frac{b-a}{4}\left[f(x_0) + 2f(x_2) + f(x_4)\right] \\
M_2 &= \sum_{i=1}^{2} f\left(\frac{x_i^L + x_i^R}{2}\right)\frac{b-a}{2} = \frac{b-a}{2}\left[f(x_1) + f(x_3)\right],
\end{aligned}$$

  In turn,

$$\begin{aligned}
\frac{1}{3}T_2 + \frac{2}{3}M_2 &= \frac{b-a}{12}\left[f(x_0) + 2f(x_2) + f(x_4)\right] + \frac{b-a}{3}\left[f(x_1) + f(x_3)\right] \\
&= \frac{b-a}{12}\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + f(x_4)\right] = S_4
\end{aligned}$$

- In **Simpson's rule**, the integrand f is approximated by a piecewise quadratic function. In comparison, in the **left point rule**, the **right point rule**, the **trapezoidal rule**, or the **midpoint rule**, the function f is approximated by a piecewise linear function.

Error bound of Simpson's rule

Theorem Suppose $f^{(4)}(x)$ is continuous and $|f^{(4)}(x)| \le K$ on $[a, b]$, $h = \frac{b-a}{n}$, then

$$\left|\int_a^b f(x)\,dx - S_n\right| \le \frac{K(b-a)^5}{180n^4} = \frac{K(b-a)}{180}h^4.$$

Example. Consider $\mathcal{I} = \int_0^1 2\cos x\ dx$, how large does n should we take so that **Simpson's rule** is accurate to within 10^{-4}.

solution.

Let $f(x) = 2\cos x$, thus, $|f^{(4)}(x)| = |2\cos x| \leq 2$, i.e., $K \leq 2$, we have

$$\frac{2}{180n^4} < 10^{-4}, \quad \Longrightarrow n > \sqrt{\frac{2}{180 \times 10^{-4}}} \approx 3.25,$$

thus, we can take $n = 4$ at least to approximate $\mathcal{I}$ within 10^{-4}.

Proof for the error bounds

- In terms of the **left point rule**, $h = \frac{b-a}{n}$, the error bound is given by

$$|E_{L_n}| = \left|\int_a^b f(x) - L_n\right| \leq \frac{K(b-a)^2}{2n} = \frac{K(b-a)}{2}h.$$

Proof. We first look at the i-th subinterval $[x_i^L, x_i^R]$, using the integration by parts, we have

$$\begin{aligned} E_i &= \int_{x_i^L}^{x_i^R} f(x)\ dx - f(x_i^L)h = \int_{x_i^L}^{x_i^R} f(x)\ d(x - x_i^R) - f(x_i^L)h \\ &= (x - x_i^R)f(x)\Big|_{x_i^L}^{x_i^R} - \int_{x_i^L}^{x_i^R} (x - x_i^R)f'(x)\ dx - f(x_i^L)h = -\int_{x_i^L}^{x_i^R} (x - x_i^R)f'(x)\ dx. \end{aligned}$$

Note that $|f'| \leq K$, we have

$$\begin{aligned} |E_i| &= \left|\int_{x_i^L}^{x_i^R} (x - x_i^R)f'(x)\ dx\right| \leq \int_{x_i^L}^{x_i^R} \left|(x - x_i^R)\right| \cdot |f'(x)|\ dx \\ &\leq K\int_{x_i^L}^{x_i^R} (x_i^R - x)\ dx = K\int_{x_i^L}^{x_i^R} (x_i^L + h - x)\ dx \\ &= K\left[hx - \frac{(x_i^L - x)^2}{2}\right]\Big|_{x_i^L}^{x_i^R} = K\left(h^2 - \frac{h^2}{2}\right) = K\frac{h^2}{2} = \frac{K(b-a)^2}{2n^2}. \end{aligned}$$

Thus,

$$|E_{L_n}| = |\sum_{i=1}^n E_i| \leq \sum_{i=1}^n |E_i| \leq \frac{K(b-a)^2}{2n^2} \cdot n = \frac{K(b-a)^2}{2n} = \frac{K(b-a)}{2}h.$$

□

- In terms of the **right point rule**, $h = \frac{b-a}{n}$, the error bound is given by

$$\begin{aligned} E_i &= \int_{x_i^L}^{x_i^R} f(x)\ dx - f(x_i^R)h = \int_{x_i^L}^{x_i^R} f(x)\ d(x - x_i^L) - f(x_i^R)h \\ &= (x - x_i^L)f(x)\Big|_{x_i^L}^{x_i^R} - \int_{x_i^L}^{x_i^R} (x - x_i^L)f'(x)\ dx - f(x_i^R)h = -\int_{x_i^L}^{x_i^R} (x - x_i^L)f'(x)\ dx. \end{aligned}$$

Note that $|f'| \leq K$, we have

$$\begin{aligned} |E_i| &= \left|\int_{x_i^L}^{x_i^R} (x - x_i^L)f'(x)\ dx\right| \leq \int_{x_i^L}^{x_i^R} \left|(x - x_i^L)\right| \cdot |f'(x)|\ dx \\ &\leq K\int_{x_i^L}^{x_i^R} (x - x_i^L)\ dx = K\int_{x_i^L}^{x_i^R} (x - x_i^R + h)\ dx \\ &= K\left[hx + \frac{(x - x_i^R)^2}{2}\right]\Big|_{x_i^L}^{x_i^R} = K\left(h^2 - \frac{h^2}{2}\right) = K\frac{h^2}{2} = \frac{K(b-a)^2}{2n^2}. \end{aligned}$$

Thus,

$$|E_{R_n}| = |\sum_{i=1}^{n} E_i| \leq \sum_{i=1}^{n} |E_i| \leq \frac{K(b-a)^2}{2n^2} \cdot n = \frac{K(b-a)^2}{2n} = \frac{K(b-a)}{2} h.$$

Exercise proof of **midpoint rule**, **trapezoidal rule** and **Simpson's rule**, (hint: using the integration by parts twice for **midpoint rule**, **trapezoidal rule** and using the integration by parts forth for **Simpson's rule**).